



 ENTERED

PCT10

RAW SEQUENCE LISTING

DATE: 02/14/2003

PATENT APPLICATION: US/10/048,146B

TIME: 13:49:16

Input Set : A:\62068.app

Output Set: N:\CRF4\02142003\J048146B.raw

3 <110> APPLICANT: Tsang et al.
 5 <120> TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING LARVAL TAENIA SOLIUM
 7 <130> FILE REFERENCE: 6395-62068
 9 <140> CURRENT APPLICATION NUMBER: 10/048,146B
 10 <141> CURRENT FILING DATE: 2000-08-03
 12 <150> PRIOR APPLICATION NUMBER: US 60/147,318
 13 <151> PRIOR FILING DATE: 1999-08-03
 15 <150> PRIOR APPLICATION NUMBER: PCT/US00/21173
 16 <151> PRIOR FILING DATE: 2000-08-03
 18 <160> NUMBER OF SEQ ID NOS: 9
 20 <170> SOFTWARE: PatentIn version 3.1
 22 <210> SEQ ID NO: 1
 23 <211> LENGTH: 2153
 24 <212> TYPE: DNA
 25 <213> ORGANISM: Taenia solium
 27 <220> FEATURE:
 28 <221> NAME/KEY: CDS
 29 <222> LOCATION: (145)..(531)
 30 <223> OTHER INFORMATION:
 W--> 33 <400> 1
 34 ctgcagtgaag gttgacaagt agttgacat ttacggaaca tcaatggagg acactttggt 60
 36 agggaaagca tacgataaac ataaaccaat gctgggtata taagagacga tctcggctac 120
 38 acttgtaact gaacaacctg taga atg cgt gcc tac att gtg ctt ctc gct 171
 39 Met Arg Ala Tyr Ile Val Leu Leu Ala
 40 1 5
 42 ctc act gtt ttc gta gtg acg gtg tcg gcc gag tgg gtg ccc att tcg 219
 43 Leu Thr Val Phe Val Val Thr Val Ser Ala Glu Trp Val Pro Ile Ser
 44 10 15 20 25
 46 agt gtc cac ata gcc tca tgc aaa agc cac tac atg ttc caa tta aaa 267
 47 Ser Val His Ile Ala Ser Cys Lys Ser His Tyr Met Phe Gln Leu Lys
 48 30 35 40
 50 cgc ttt ttt gcc ttt agg aaa aac aaa ccg aaa gat gtt gca aat agt 315
 51 Arg Phe Phe Ala Phe Arg Lys Asn Lys Pro Lys Asp Val Ala Asn Ser
 52 45 50 55
 54 acg aaa aaa ggg ata gaa tat gtc cac gaa ttc ttc cac gaa gac ccg 363
 55 Thr Lys Lys Gly Ile Glu Tyr Val His Glu Phe Phe His Glu Asp Pro
 56 60 65 70
 58 att ggt aaa caa att gct caa ctc gca aag gaa tgg aag gaa gca atg 411
 59 Ile Gly Lys Gln Ile Ala Gln Leu Ala Lys Glu Trp Lys Glu Ala Met
 60 75 80 85
 62 ttg gaa ggt agg ttt tgg tgt ttt ctg tca gaa gaa aat tat cta ttc 459
 63 Leu Glu Gly Arg Phe Trp Cys Phe Leu Ser Glu Glu Asn Tyr Leu Phe
 64 90 95 100 105



RAW SEQUENCE LISTING

DATE: 02/14/2003

PATENT APPLICATION: US/10/048,146B

TIME: 13:49:16

Input Set : A:\62068.app

Output Set: N:\CRF4\02142003\J048146B.raw

```

66 att cat cta gac aaa ggc aaa ata cgg acg tca ctg gtt gag cac tgc      507
67 Ile His Leu Asp Lys Gly Lys Ile Arg Thr Ser Leu Val Glu His Cys
68              110              115              120
70 aaa ggt cct aag aaa aaa act gct taacttgtca actttcatgc gttcttctct      561
71 Lys Gly Pro Lys Lys Lys Thr Ala
72              125
74 tcactaataa atgctcatta ataagaaagc tgccttttgc aagatcaacg agggccatag      621
76 actgtgaggg ttatagccta aggttatggg gtgaaatgag ataggaattg agcatttgag      681
78 aagttactaa tttaaattga aagccgcatt tcttctgcaa ttgacgtgtg atggttagcg      741
80 aaaccaagtg aagcacgacc tcttgagtcg tttcaacagc cgccagtggg ttcaccagtg      801
82 gcttcaccag tgggtagact ggtttgtcac acatgcgagg tacggtcaga gggctaacag      861
84 gtgtggtgga ggggccaaca cgtgtaagac aagcagttcc ccttctctgt cgtgaggcac      921
86 actcagcacc cacctcgttt acttctccct tgacgactgt aatgcatttg gggtcacccat      981
88 gcccccgcca agttgaaggc actgatgaca tttgtacccat atcaccgata agtattaact     1041
90 cttccacttc ccagattttg aggtcaggcg atcctactga ctcggtgtag ccccatgggtg     1101
92 gtccatgctc tgcaccattc gctgttcagt ggagcatcca cctagacggc caaccaatct     1161
94 cgctccctt ctcctgtgct caagatgtgc gtcggtgaga tttggagggt ctgatcacca     1221
96 tactaaccac gtaggtttca tcatctctaa gaagcaccac ttcttgagggt cgcattgtgt     1281
98 accaccagcc ggtgtaatca agagtgactt tcgcgtcacc cctaagaagg ctatagatct     1341
100 gcaagtcagc gcaatagctt cagccatgct gactaaaatg tgtaagggac cagtagctct     1401
102 agcccaacac aagtggagct aataatgggc ttccccagat acatgaatcc caaatcggtg     1461
104 agcatggggc atgaatatgg ccttctgagt cttccttgaa tgcaaacgaa ggcatagcac     1521
106 gagggtagga tgagtgtaca gaaaacagcg aggcaacgaa tctactggca tggccctgat     1581
108 gccaccccgcc ccagctaggg tagtttggcc acctcagtcc ttaatcgaat gcggcagtca     1641
110 gaacaaacaa agtattacat agccacactc ttcttttgag cgtcgtcctc gacgctcctt     1701
112 tcgacacacc tcccgcacat gccaccacaa agtaatcagt actggggaga caccacagag     1761
114 ctaaccgtgc cagtcacatga aaatttgacg gcaactgagg agatgcctga ccccttttg     1821
116 cagttcgaat gctgcccggt gtcacactcc tgcatcagcc atcacctacg attcaaacat     1881
118 cctagtcgcc aaattttcgt gaacctctta aaattttcgt gcactctcaa gacacttcca     1941
120 actgacttag agctttttca tttggtgaga acacgtaaaa gcttcaagta aacaacaggc     2001
122 aacgatttca ctttgatgct ctacccatca attctcttgt atgtgccacc accttaaacc     2061
124 ctccctgacc acttccactc tctctctctc cctaaataac aacacttggg agcatgaatg     2121
126 gtgtctgtca aagttacacc cctagactgc ag                                2153
129 <210> SEQ ID NO: 2
130 <211> LENGTH: 129
131 <212> TYPE: PRT
132 <213> ORGANISM: Taenia solium
134 <400> SEQUENCE: 2
136 Met Arg Ala Tyr Ile Val Leu Leu Ala Leu Thr Val Phe Val Val Thr
137 1              5              10              15
140 Val Ser Ala Glu Trp Val Pro Ile Ser Ser Val His Ile Ala Ser Cys
141              20              25              30
144 Lys Ser His Tyr Met Phe Gln Leu Lys Arg Phe Phe Ala Phe Arg Lys
145              35              40              45
148 Asn Lys Pro Lys Asp Val Ala Asn Ser Thr Lys Lys Gly Ile Glu Tyr
149              50              55              60
152 Val His Glu Phe Phe His Glu Asp Pro Ile Gly Lys Gln Ile Ala Gln
153 65              70              75              80
156 Leu Ala Lys Glu Trp Lys Glu Ala Met Leu Glu Gly Arg Phe Trp Cys

```

RAW SEQUENCE LISTING

DATE: 02/14/2003

PATENT APPLICATION: US/10/048,146B

TIME: 13:49:16

Input Set : A:\62068.app

Output Set: N:\CRF4\02142003\J048146B.raw

```

157          85          90          95
160 Phe Leu Ser Glu Glu Asn Tyr Leu Phe Ile His Leu Asp Lys Gly Lys
161          100          105          110
164 Ile Arg Thr Ser Leu Val Glu His Cys Lys Gly Pro Lys Lys Lys Thr
165          115          120          125
168 Ala
172 <210> SEQ ID NO: 3
173 <211> LENGTH: 298
174 <212> TYPE: DNA
175 <213> ORGANISM: Taenia solium
177 <220> FEATURE:
178 <221> NAME/KEY: CDS
179 <222> LOCATION: (3)..(224)
180 <223> OTHER INFORMATION:
W--> 183 <400> 3
184 ta ttc gta gtg gcg gtt tcg gcc gag aaa aac aaa ccg aag tgt gat      47
185   Phe Val Val Ala Val Ser Ala Glu Lys Asn Lys Pro Lys Cys Asp
186   1          5          10          15
188 gca aat agt act aag aaa gag ata gaa tat atc cac aat tgg ttt ttc      95
189 Ala Asn Ser Thr Lys Lys Glu Ile Glu Tyr Ile His Asn Trp Phe Phe
190          20          25          30
192 cat gat gac ccg att gga aaa caa att gct caa ctc gca aag gac tgg      143
193 His Asp Asp Pro Ile Gly Lys Gln Ile Ala Gln Leu Ala Lys Asp Trp
194          35          40          45
196 aat gaa aca gtg cag gaa gcc aaa ggc aaa ttt tgg gcg tca ctg gct      191
197 Asn Glu Thr Val Gln Glu Ala Lys Gly Lys Phe Trp Ala Ser Leu Ala
198          50          55          60
200 gag tac tgc aga ggt ctg aag aac aaa act gct taacttgtca actttcatgc      244
201 Glu Tyr Cys Arg Gly Leu Lys Asn Lys Thr Ala
202          65          70
204 gttctttctct tcaccaataa atgctgatta acaagaaaaa aaaaaaaaaa aaaa      298
207 <210> SEQ ID NO: 4
208 <211> LENGTH: 74
209 <212> TYPE: PRT
210 <213> ORGANISM: Taenia solium
212 <400> SEQUENCE: 4
214 Phe Val Val Ala Val Ser Ala Glu Lys Asn Lys Pro Lys Cys Asp Ala
215 1          5          10          15
218 Asn Ser Thr Lys Lys Glu Ile Glu Tyr Ile His Asn Trp Phe Phe His
219          20          25          30
222 Asp Asp Pro Ile Gly Lys Gln Ile Ala Gln Leu Ala Lys Asp Trp Asn
223          35          40          45
226 Glu Thr Val Gln Glu Ala Lys Gly Lys Phe Trp Ala Ser Leu Ala Glu
227          50          55          60
230 Tyr Cys Arg Gly Leu Lys Asn Lys Thr Ala
231 65          70
234 <210> SEQ ID NO: 5
235 <211> LENGTH: 294
236 <212> TYPE: DNA

```

RAW SEQUENCE LISTING

DATE: 02/14/2003

PATENT APPLICATION: US/10/048,146B

TIME: 13:49:16

Input Set : A:\62068.app

Output Set: N:\CRF4\02142003\J048146B.raw

237 <213> ORGANISM: Taenia solium
 239 <220> FEATURE:
 240 <221> NAME/KEY: CDS
 241 <222> LOCATION: (3)..(221)
 242 <223> OTHER INFORMATION:
 W--> 245 <400> 5
 246 tt ttc gta gtg gcg gtg tcg gcc gag gaa act aaa cca gag gac gtg 47
 247 Phe Val Val Ala Val Ser Ala Glu Glu Thr Lys Pro Glu Asp Val
 248 1 5 10 15
 250 gta aag aat att aag aaa ggg atg gaa gtt gtc tac aaa ttt ttc tac 95
 251 Val Lys Asn Ile Lys Lys Gly Met Glu Val Val Tyr Lys Phe Phe Tyr
 252 20 25 30
 254 gaa gac ccg ttg gga aag aaa ata gct caa ctc gca aag gac tgg aag 143
 255 Glu Asp Pro Leu Gly Lys Lys Ile Ala Gln Leu Ala Lys Asp Trp Lys
 256 35 40 45
 258 gaa gca atg ttg gaa gcc aga agc aaa gtg cgg gcg tca ctg gct gag 191
 259 Glu Ala Met Leu Glu Ala Arg Ser Lys Val Arg Ala Ser Leu Ala Glu
 260 50 55 60
 262 tac atc aga ggt ctc aag aac gaa gct gct taacttgtca actttcatgc 241
 263 Tyr Ile Arg Gly Leu Lys Asn Glu Ala Ala
 264 65 70
 266 gttcttctct tcactaataa atgctcatta ataagaaaaa aaaaaaaaaa aaa 294
 269 <210> SEQ ID NO: 6
 270 <211> LENGTH: 73
 271 <212> TYPE: PRT
 272 <213> ORGANISM: Taenia solium
 274 <400> SEQUENCE: 6
 276 Phe Val Val Ala Val Ser Ala Glu Glu Thr Lys Pro Glu Asp Val Val
 277 1 5 10 15
 280 Lys Asn Ile Lys Lys Gly Met Glu Val Val Tyr Lys Phe Phe Tyr Glu
 281 20 25 30
 284 Asp Pro Leu Gly Lys Lys Ile Ala Gln Leu Ala Lys Asp Trp Lys Glu
 285 35 40 45
 288 Ala Met Leu Glu Ala Arg Ser Lys Val Arg Ala Ser Leu Ala Glu Tyr
 289 50 55 60
 292 Ile Arg Gly Leu Lys Asn Glu Ala Ala
 293 65 70
 296 <210> SEQ ID NO: 7
 297 <211> LENGTH: 6
 298 <212> TYPE: PRT
 299 <213> ORGANISM: Taenia solium
 301 <400> SEQUENCE: 7
 303 Ile Ala Gln Leu Ala Lys
 304 1 5
 307 <210> SEQ ID NO: 8
 308 <211> LENGTH: 24
 309 <212> TYPE: PRT
 310 <213> ORGANISM: Taenia solium
 312 <220> FEATURE:

RAW SEQUENCE LISTING

DATE: 02/14/2003

PATENT APPLICATION: US/10/048,146B

TIME: 13:49:16

Input Set : A:\62068.app

Output Set: N:\CRF4\02142003\J048146B.raw

313 <221> NAME/KEY: variant
 314 <222> LOCATION: (7)..(8)
 315 <223> OTHER INFORMATION: Amino acid at position 7 may also be valine
 318 <220> FEATURE:
 319 <221> NAME/KEY: site
 320 <222> LOCATION: (21)..(22)
 321 <223> OTHER INFORMATION: Asparagine at position 21 is an amino acid insertion
 324 <220> FEATURE:
 325 <221> NAME/KEY: variant
 326 <222> LOCATION: (14)..(15)
 327 <223> OTHER INFORMATION: Amino acid at position 14 may also be glycine
 330 <220> FEATURE:
 331 <221> NAME/KEY: variant
 332 <222> LOCATION: (18)..(19)
 333 <223> OTHER INFORMATION: Amino acid at position 18 may also be valine
 336 <220> FEATURE:
 337 <221> NAME/KEY: variant
 338 <222> LOCATION: (19)..(20)
 339 <223> OTHER INFORMATION: Amino acid at position 19 may also be histidine
 342 <220> FEATURE:
 343 <221> NAME/KEY: variant
 344 <222> LOCATION: (20)..(21)
 345 <223> OTHER INFORMATION: Amino acid at position 20 may also be arginine
 348 <400> SEQUENCE: 8
 350 Lys Asn Lys Pro Lys Asp Asp Ala Ala Ser Thr Lys Lys Glu Ile Glu
 351 1 5 10 15
 354 Tyr Ile Trp His Asn Phe Phe Phe
 355 20
 358 <210> SEQ ID NO: 9
 359 <211> LENGTH: 13
 360 <212> TYPE: PRT
 361 <213> ORGANISM: Taenia solium
 363 <220> FEATURE:
 364 <221> NAME/KEY: variant
 365 <222> LOCATION: (5)..(6)
 366 <223> OTHER INFORMATION: Amino acid at position 5 may also be isoleucine
 369 <220> FEATURE:
 370 <221> NAME/KEY: variant
 371 <222> LOCATION: (12)..(13)
 372 <223> OTHER INFORMATION: Amino acid at position 12 may also be aspartic acid
 375 <220> FEATURE:
 376 <221> NAME/KEY: variant
 377 <222> LOCATION: (7)..(9)
 378 <223> OTHER INFORMATION: Amino acid at position 7 may also be asparagine
 381 <220> FEATURE:
 382 <221> NAME/KEY: site
 383 <222> LOCATION: (8)..(9)
 384 <223> OTHER INFORMATION: Tryptophan at position 8 is an amino acid insertion
 387 <400> SEQUENCE: 9

VERIFICATION SUMMARY

DATE: 02/14/2003

PATENT APPLICATION: US/10/048,146B

TIME: 13:49:17

Input Set : A:\62068.app

Output Set: N:\CRF4\02142003\J048146B.raw

L:33 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:1,Line#:30
L:183 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:3,Line#:180
L:245 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:5,Line#:242